Cross-lingual Projection of Role-Semantic Information

Sebastian Padó

Computational Linguistics
Saarland University
pado@coli.uni-sb.de

October 28, 2005
1 Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck

2 Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection

3 Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models
Outline

1 Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck

2 Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection

3 Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models
The task of automatically identifying the **semantic roles** conveyed by sentential constituents.
Shallow Semantic Parsing

The task of automatically identifying the **semantic roles** conveyed by sentential constituents.

- Relevant for several applications (IE, IR, QA)
- Common semantic representation across languages
Frame Semantics

Role-semantics paradigm based on **conceptual** structures (Filmore et al., 2003).

<table>
<thead>
<tr>
<th>Frame Elements</th>
<th>FEEs</th>
</tr>
</thead>
</table>
| **Cognizer**   | Peter knows the situation.  
Pat believes that things will change. |
| **Content**    | Peter knows **the situation**.  
Pat believes **that things will change**. |
| **FEEs**       | aware.v, believe.v, comprehend.v, conceive.v, imagine.v, know.v, belief.n, consciousness.v, hunch.n, suspicion.v, conscious.a, knowledgeable.a |
1 Motivation
- Shallow Semantic Parsing
- Knowledge Acquisition Bottleneck

2 Role Projection in a Parallel Corpus
- Word-based Projection
- Syntax-based Projection

3 Projection Results
- Experimental Set-up
- Evaluation of Projection Models
Knowledge Acquisition Bottleneck

Data-driven development of shallow semantic parsers (see e.g. Carreras and Màrquez, 2005) requires:

1. English FrameNet lexicon (> 500 frames, > 7,000 lemmas)
2. English annotated example sentences (100,000 available)
Knowledge Acquisition Bottleneck

- Data-driven development of shallow semantic parsers (see e.g. Carreras and Màrquez, 2005) requires:
  1. English FrameNet lexicon (> 500 frames, > 7,000 lemmas)
  2. English annotated example sentences (100,000 available)

- Frame Semantics is (largely) language-independent: annotation efforts for German, Spanish, and Japanese

- Annotation laborious and time-consuming
Knowledge Acquisition Bottleneck

- Data-driven development of shallow semantic parsers (see e.g. Carreras and Màrquez, 2005) requires:
  1. English FrameNet lexicon (> 500 frames, > 7,000 lemmas)
  2. English annotated example sentences (100,000 available)

- Frame Semantics is (largely) language-independent:
  annotation efforts for German, Spanish, and Japanese

- Annotation laborious and time-consuming

**Knowledge Acquisition Bottleneck:**
Can we reduce annotation effort for new languages?
Main Ideas

- Use English FrameNet resource as basis
- Project information to other languages using parallel corpora

Two steps:
1. Project FrameNet lexicon (IGK meeting in Mertesdorf)
2. Project role information (now)
Outline

1 Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck

2 Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection

3 Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment

Peter knows the situation.

Peter kennt die Situation.
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence

- Awareness
- Cognizer
- Content

Peter knows the situation.

Peter kennt die Situation.
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame

![Diagram of role projection](attachment:image.png)

- **Awareness**
- **Cognizer**
- **Content**

Peter knows the situation.
Peter kennt die Situation.
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame
4. Project roles from source to target sentence

---

```
1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame
4. Project roles from source to target sentence
```
Role Projection in a Parallel Corpus

Start with bi-sentence (translation) with word alignment
Obtain role-semantic analysis for source sentence
Check if target predicate can evoke the same frame
Project roles from source to target sentence

Assumption:
Bi-sentences have parallel (role) semantics
Role Projection in a Parallel Corpus

1. Start with bi-sentence (translation) with word alignment
2. Obtain role-semantic analysis for source sentence
3. Check if target predicate can evoke the same frame
4. Project roles from source to target sentence

Assumption:
Bi-sentences have parallel (role) semantics

Empirical result:
For English / German, 92% of roles match
Related Work

- Induction of multilingual morphological analyzers (Mann and Yarowsky, 2001)
- Projection of POS-tag information (Yarowsky et al., 2001)
- Projection of bracketing information (Yarowsky et al., 2001)
- Projection of dependency relations (Hwa et al., 2002)
Outline

1 Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck

2 Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection

3 Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models
Word-based Projection

1. For each source semantic role, identify word span

For example:

- "Peter and Mary left." to "Peter und auch Maria gingen."
- The theme "Departing" is projected.

Sebastian Padó  Cross-lingual Projection of Role-Semantic Information  21
Word-based Projection

1. For each source semantic role, identify word span
2. Follow all word alignments

Motivation
Role Projection in a Parallel Corpus
Projection Results
Conclusions
Word-based Projection

1. For each source semantic role, identify word span
2. Follow all word alignments
3. Target role span is union of all projections

Example:

- **Peter and Mary** left.
- **Peter und auch Maria** gingen.

Context:

Motivation
Role Projection in a Parallel Corpus
Projection Results
Conclusions

Word-based Projection

Syntax-based Projection

Sebastian Padó
Cross-lingual Projection of Role-Semantic Information
Word-based Projection

1. For each source semantic role, identify word span
2. Follow all word alignments
3. Target role span is union of all projections

Missing word alignments: convex complementing heuristic
Word-based Projection

1. For each source semantic role, identify word span
2. Follow all word alignments
3. Target role span is union of all projections

```
Peter and Mary left.
Peter und auch Maria gingen.
```

Missing word alignments: convex complementing heuristic
Outline

1. Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck

2. Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection

3. Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models
For each source role, identify source constituent(s)

1. For each source role, identify source constituent(s)

2. Find optimal alignment between S and T constituents

3. Label target constituent(s) with role

Peter and Mary left.
Peter und auch Maria gingen.

Departing
Theme
NP Peter and Mary left.
Peter und auch Maria gingen.
Departing
Syntax-based Projection

1. For each source role, identify source constituent(s)
2. Find **optimal alignment** between S and T constituents
Syntax-based Projection

1. For each source role, identify source constituent(s)
2. Find **optimal alignment** between S and T constituents
3. Label target constituent(s) with role

For example:

**Word-based Projection**

- For each source role, identify source constituent(s)
- Find optimal alignment between S and T constituents
- Label target constituent(s) with role

**Syntactic Projection**

- For each source role, identify source constituent(s)
- Find optimal alignment between S and T constituents
- Label target constituent(s) with role

---

For the example:

- **NP** Peter and Mary left.
- **NP** Peter und auch Maria gingen.

**Departing**

**Theme**

Sebastian Padó

Cross-lingual Projection of Role-Semantic Information 29
Syntax-based Projection

1. For each source role, identify source constituent(s)
2. Find optimal alignment between S and T constituents
3. Label target constituent(s) with role

Can use any bracketing information (chunks, “full” constituents)
Probabilistic Constituent Alignment

- Two sets of constituents, $C$ and $C'$
- For each $c \in C$, find $c' \in C'$ with maximal word overlap
Probabilistic Constituent Alignment

- Two sets of constituents, $C$ and $C'$
- For each $c \in C$, find $c' \in C'$ with maximal word overlap

**Forward alignment**
- Align from source to target constituents
- Assumes one target constituent per source constituent

![Diagram](attachment:image_url)
Probabilistic Constituent Alignment

- Two sets of constituents, $C$ and $C'$
- For each $c \in C$, find $c' \in C'$ with maximal word overlap

1. **Forward** alignment
   - Align from source to target constituents
   - Assumes one target constituent per source constituent

2. **Backward** alignment
   - Aligns from target to source constituents
   - Source constituents can correspond to none or $>1$ target constituents
Outline

1 Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck

2 Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection

3 Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models
Experimental Set-up

Data
- Sample of 1000 English-German Bi-sentences from EUROPARL (Koehn, 2000)
- Choice informed by FrameNet (E) and SALSA (D) lexicons
- Two sides of each bi-sentence annotated independently
  - Annotators tagged equal amount of English and German
  - Inter-annotator agreement: $\kappa = 0.84$
- Word alignment: GIZA++ (Och and Ney, 2003)
## Experimental Set-up

### Method

- Project roles from English gold annotation onto German
- Evaluate against German gold annotation
- Compare word-based, chunk-based, and constituent-based models
  - Chunk-based models use Abney’s (1997, E) and Schmid and Schulte im Walde’s (2000, D) base NP chunkers
  - Constituent-based models use Collins’ (1997, E) and Dubey’s (2003, D) parsers
Outline

1 Motivation
   - Shallow Semantic Parsing
   - Knowledge Acquisition Bottleneck

2 Role Projection in a Parallel Corpus
   - Word-based Projection
   - Syntax-based Projection

3 Projection Results
   - Experimental Set-up
   - Evaluation of Projection Models
Word-based Projection

<table>
<thead>
<tr>
<th>Model</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>WordAlign</td>
<td>0.41</td>
<td>0.40</td>
<td>0.41</td>
</tr>
<tr>
<td>WordAlign + ConvexComp</td>
<td>0.46</td>
<td>0.45</td>
<td>0.46</td>
</tr>
<tr>
<td>UpperBnd</td>
<td>0.85</td>
<td>0.84</td>
<td>0.84</td>
</tr>
</tbody>
</table>

- Alignments can guide projection task substantially
- WordAlign exploits no linguistic information
- ConvexComp improves the F-score
- F-score sig worse than UpperBnd
### Chunk-based Projection

<table>
<thead>
<tr>
<th>Model</th>
<th>Prec</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>WordAlign + ConvexComp</td>
<td>0.46</td>
<td>0.45</td>
<td>0.46</td>
</tr>
<tr>
<td>ForwardAlign</td>
<td>0.46</td>
<td>0.25</td>
<td>0.32</td>
</tr>
<tr>
<td>BackwardAlign</td>
<td>0.30</td>
<td>0.24</td>
<td>0.27</td>
</tr>
<tr>
<td>BackwardAlign + ConvexComp</td>
<td>0.32</td>
<td>0.26</td>
<td>0.29</td>
</tr>
<tr>
<td>UpperBnd</td>
<td>0.85</td>
<td>0.84</td>
<td>0.84</td>
</tr>
</tbody>
</table>

- ForwardAlign sig worse Recall than WordAlign
- F-score sig worse than UpperBnd
- Problem: Often, no chunks for source **and** target role span
  - Overlap maximisation does not yield sensible results
## Syntax-based Projection

<table>
<thead>
<tr>
<th>Model</th>
<th>Prec.</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>WordAlign + ConvexComp</td>
<td>0.46</td>
<td>0.45</td>
<td>0.46</td>
</tr>
<tr>
<td>ForwardAlign</td>
<td>0.70</td>
<td>0.60</td>
<td>0.65</td>
</tr>
<tr>
<td>BackwardAlign</td>
<td>0.60</td>
<td>0.46</td>
<td>0.52</td>
</tr>
<tr>
<td>BackwardAlign + ConvexComp</td>
<td>0.74</td>
<td>0.56</td>
<td>0.64</td>
</tr>
<tr>
<td>UpperBnd</td>
<td>0.85</td>
<td>0.84</td>
<td>0.84</td>
</tr>
</tbody>
</table>

- ForwardAlign sig better than WordAlign and BackwardAlign; sig worse than UpperBnd
- One-to-one assumption (ForwardAlign) mostly warranted
- Less guided alignment (BackwardAlign) requires ConvexComp
Error Analysis

1. Wrong or missing word alignments (e.g., PPs)
   - He asks [MSG for a doctor].
   - Er fragt nach einem Arzt.

2. Translational divergences (wrong or missing projections)
   - We claim and [SPKR we] say [MSG ...]
   - [SPKR Wir] behaupten und -- sagen [MSG ...]
Conclusions

Summary

- Principled framework for role projection
- Semantic roles can be projected between languages
- Bracketing can make up for problems in word alignment
- Best model performs at 0.65 F-Score (UpperBnd is 0.84)
- Base NP chunks not sufficient